A Network Model of Memory System Composed of Short-Term Memory and Long-Term Memory

長期記憶と閉回路をなす短期記憶のモデル

A Network Model of Memory System Composed of Short-Term Memory and Long-Term Memory

神奈川大学、工、物理、宇佐見義之

Institute of Physics, Kanagawa University, Rokkakubashi 3-27-1, Kanagawa-ku, Yokohama 221

Abstract

We study about a mathematical network model composed of short-term memory and long-term memory in the framework of neural network theory. Long-term memory area in the model plays a role to store associative memories for a long time. Short-term memory helps to learn new things and store it in long-term memory area by forming closed circuit with long-term memory. The proposed mathematical network model is constructed based on the closed circuit model founded by many physiologists. We show numerical simulation results performed upon this model and discuss the correspondences with physiological findings.

1. Introduction

It is well known that hippocampus and amygdala play an important role to learn new things and store it in the memory system. If these areas are once destroyed by some reason, the person or the animal shows the behavior that who can not remember new things. It is also believed, however, that these are not the areas to keep memories in it. In this sense we call these areas as short-term memory (STM) in this paper. Physiologists Mishikin[1], Iwai[2] and the others stated that closed circuit between hippocampus/amygdala and area TE might be formed during the process of visual recognition and learning. Area TE is considered a place storing visual memories for a long time, hence we call such area as long-term memory (LTM) area. Although many other areas are considered to contribute to visual recognition and learning, we propose a simplest version of neural network model composed of STM and LTM, which has two abilities in the same architecture. Short-term memory in the model plays a role for circulating pulses between long-term memory if there is associative input besides main input. During the pulses circulating in closed circuit, plastic synapse in LTM is strengthened, namely, learning is accomplished. The system without STM loses the ability of learning, however, LTM can recognize things that already learn.

On mathematical viewpoint the proposed model is based on the competitive and recurrent neural network model formulated by Amari[3] and the others[4]. Doubly stable potential function of plastic synapse proposed earlier by the author[5] is introduced to synaptic connection in long-term memory area. Memories are stored very stably in this area, because synaptic weight has two stable states: 0 no-connection and 1 full-connection.

2. Structure of the neural system composed of LTM and STM

Figure 1 shows architecture of the system composed of STM and STM. We consider about higher functional of visual system, so we hypothesize that LTM corresponds to area TE and STM corresponds to hippocampus or amygdala in this article. Cells x and y are input cell and output cell in LTM, respectively, connected with plastic connection W(t). There are fixed value synaptic connections between v and u as ω_{21} and between v and z as ω_{22} in STM area.Total response of the system is given as f(y), where f is the step function. For simplicity we treat only on one connection in LTM here, however, discussion can be naturally extended to many connections of plastic synapse as usual neural network models.

We will hereafter call the name of cells and electric potential of it as the same one, because there are one to one correspondences. Now we write the equation of motion describing temporal behavior of the proposed system:

$$\begin{aligned}
\frac{dx}{dt} &= -kx + \omega_1 f(v) + I(t) - h_x, \\
\frac{dy}{dt} &= -ky + \omega_1 W f(x) + J(t) - h_y, \\
\frac{du}{dt} &= -ku + \omega_1 f(y) - h_u, \\
\frac{dv}{dt} &= -kv + \omega_{21} f(u) + \omega_{22} f(z) - h_v, \\
\frac{dz}{dt} &= -kz + J(t) - h_z,
\end{aligned} \tag{1}$$

where x,y,$\cdots$,z represent electric potential of each cell. Parameters ω_1, ω_{21}, ω_{22} and k represent value of synaptic connections indicated in Fig.1 and decay constant, respectively. I(t) and J(t) are inputs to the system as mentioned above.

We write time evolution of plastic synaptic weight in LTM as

$$\frac{dW}{dt} = -\frac{d\Phi(W)}{dW} + \alpha f(x) f(y). \tag{2}$$

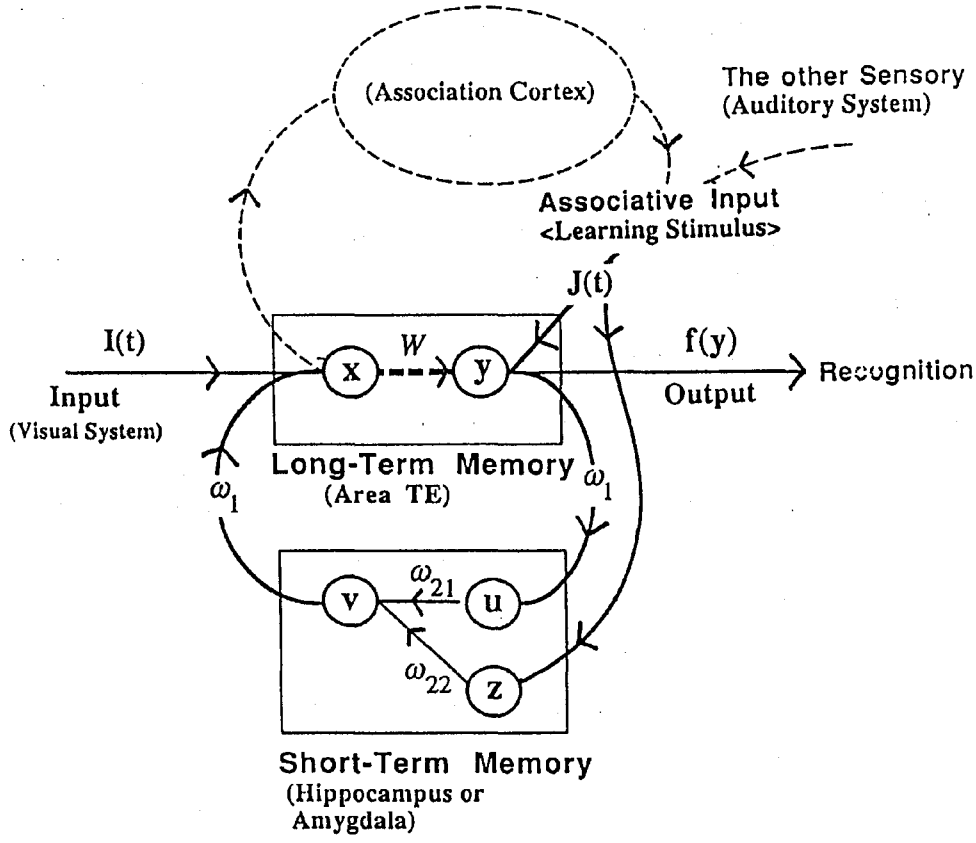


Fig.1 The structure of the proposed neural network model composed of long-term memory and short-term memory. Plastic synapse W can be strengthened by the formation of closed circuit between them, if associative input is coming to y cell in LTM and to z cell in STM.

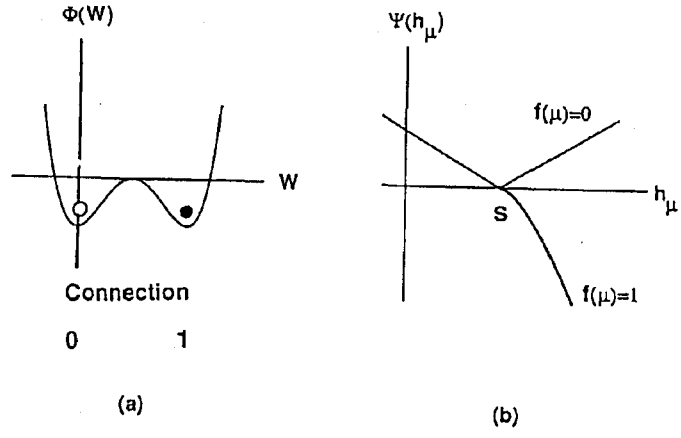


Fig.2 (a): Doubly stable potential as a function of synaptic weight W in LTM. Stable states $W = 0$ and $W = 1$ correspond to no-connection and full-connection, respectively. (b) Potential function describing fatigue effect of active cell. If a cell is firing i.e. $f(\mu) = 1$, threshold value h_μ will increase, where μ stands for each cell, $\mu = x, y, \cdot, z$. While for the state of $f(\mu) = 0$, threshold value will converge to a certain constant s.

Weight is strengthened by Hebb rule by constant factor α. And we assume that synaptic connection has two stable states, $W = 0$ no-connection and $W = 1$ full-connection, which is realized by introducing doubly stable potential function $\Phi(W)$:

$$\Phi(W) = -(W - \frac{1}{2})^2 + 2(W - \frac{1}{2})^4. \tag{3}$$

This potential is shown in Fig.2(a). We further introduce fatigue effect for cell activity of as rising threshold value:

$$\frac{dh_\mu}{dt} = -\frac{d\Psi(h_\mu)}{dh_\mu}, \tag{4}$$

$$\Psi(h_\mu) = \begin{cases} -\varepsilon_1 \ (h_\mu - s)^{\frac{3}{2}} & \text{for} \quad f(\mu) = 1 \\ \varepsilon_2 \ sgn(h_\mu - s) \ h_\mu & \text{for} \quad f(\mu) = 0 \end{cases} \tag{5}$$

where $\mu(t)$ stands for each potential, $\mu(t) = x(t), y(t), \cdots, z(t)$. Hence $h_\mu(t)$ express $h_x(t), h_y(t), \cdots, h$ respectively. Parameters ε_1 and ε_2 are constants. We show the potential $\Phi(h_\mu)$ in Fig.2(b). Threshold h_μ increases when $f(\mu) = 1$ i.e. the cell is firing. And when threshold value exceeds electric potential, h_μ will converge to a certain constant value s. According to this effect, firing cell eventually cease its activity.

Finally we note the parameters used in the numerical simulation: $\omega_1 = 8, \omega_{21} = .7, \omega_{22} = 4, \alpha = .8, s = .3, \varepsilon_1 = 1.5, \varepsilon_2 = 1.2$

3. Temporal behavior of the system

We summarize the temporal behavior of the proposed system in Fig.3. Upper half shows active cells and its graphical explanation. Lower half shows time evolution of every time-dependent variables. We give the following three inputs to the system with a certain strength $a = 1.3$: (1) At T1, main input to x.

(2) At T2, input to x, and associative input to y and u.

(3) At T4, main input to x.

At time T1, cell x receives input I(t) as $I(t) = a\delta(t - T1)$. Because the initial value of synaptic weight W is set as zero, signal is not transmitted to output cell y. This means system does not recognize the input signal. At time T2, we give the associative input besides main input. Then around time T3 region shown in Fig.3, we can observe the formation of closed circuit between LTM and STM. During this period, plastic synapse W is strengthened by Hebb rule, namely, W increases when x and y are both activated. The bottom graph in

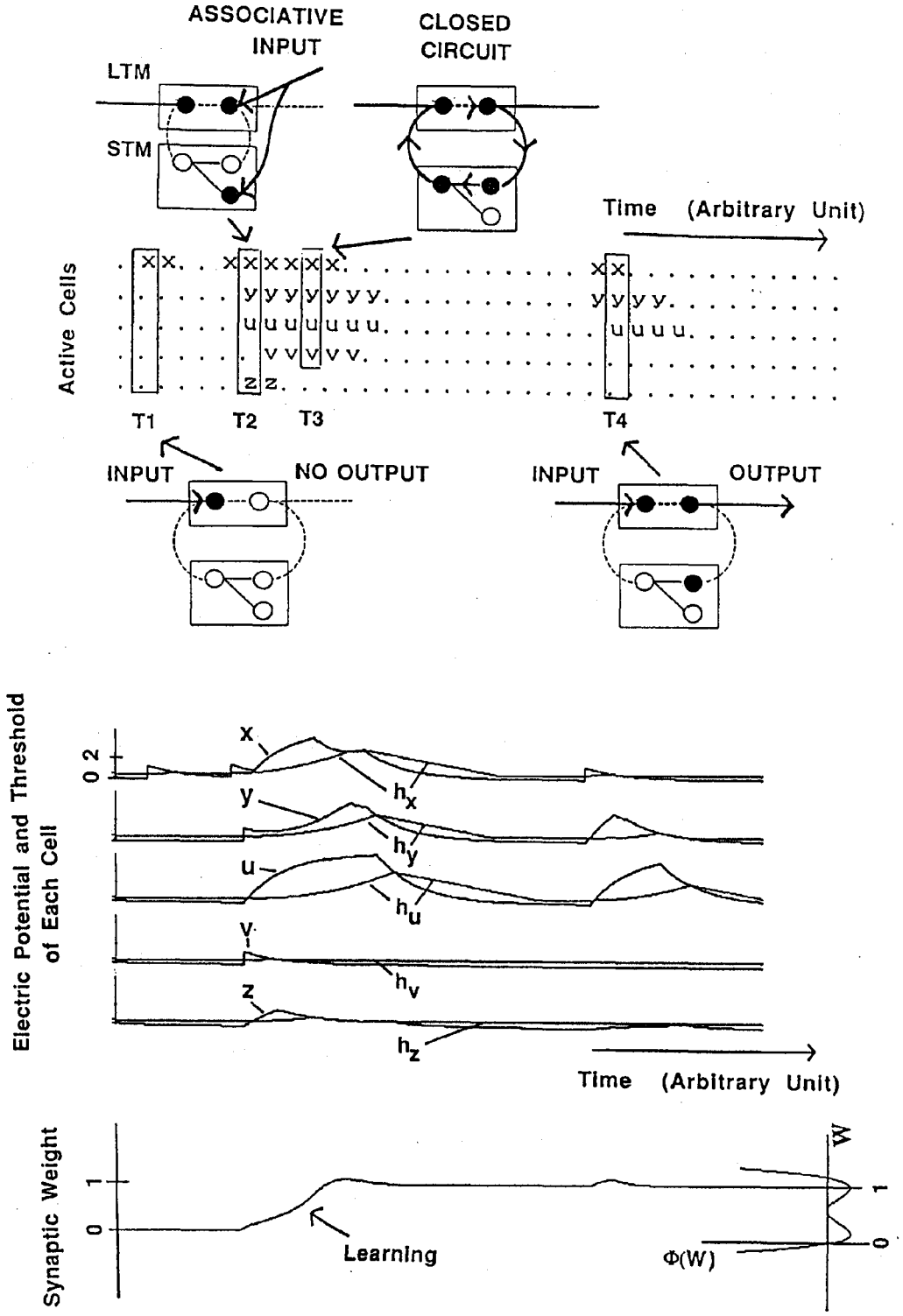


Fig.3 Simulation results with its graphical explanation. The upper half shows active cells, namely each character expresses firing cell. Stimulus are given at T1, T2 and T4. The lower half shows time evolution of electric potential,threshold value of each cell and synaptic weight. We can observe the growth of synaptic weight. It means that learning has been accomplished in LTM area during closed circuit was formed.

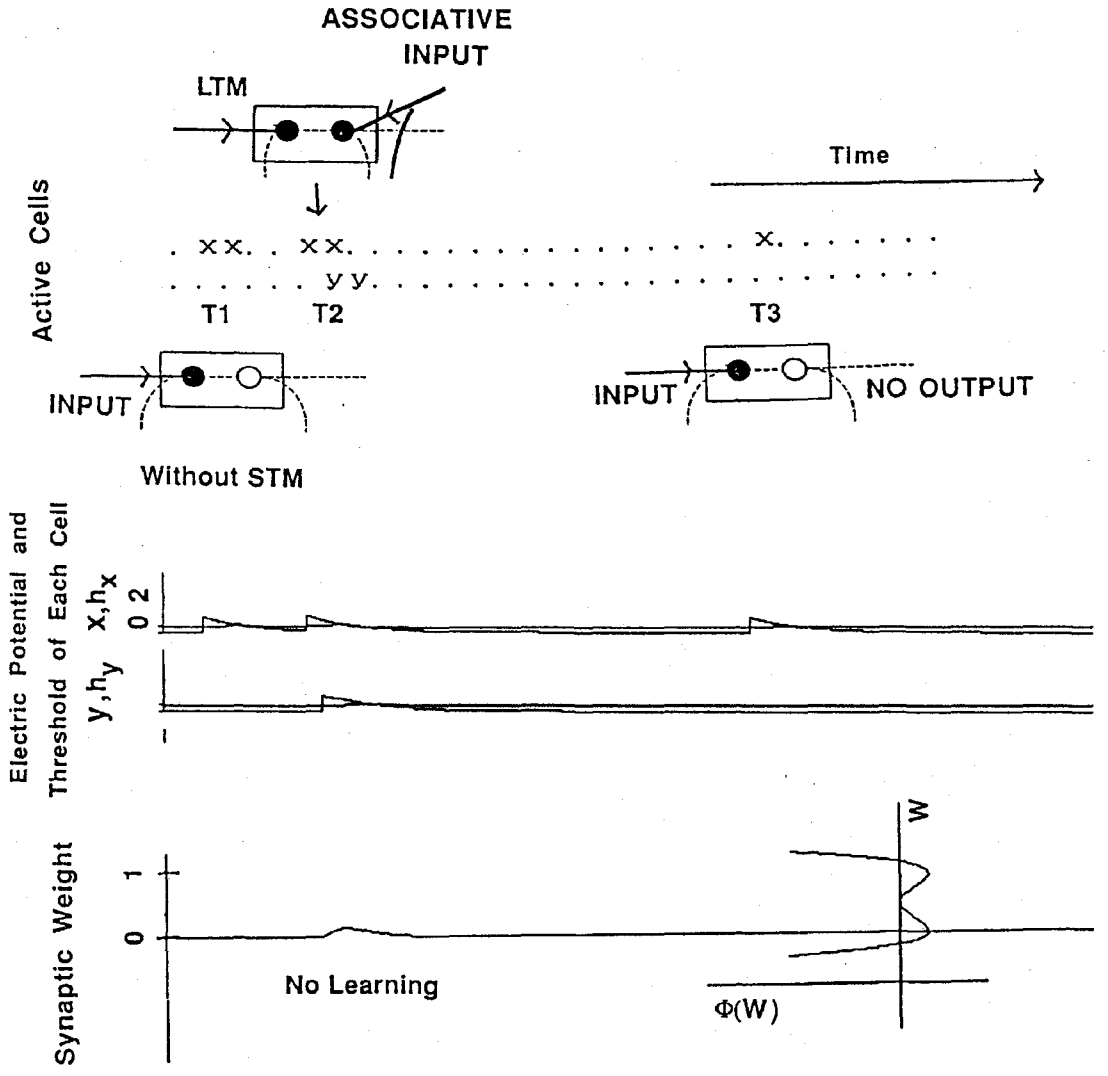


Fig.4 Temporal behavior of the system without short-term memory. The same inputs are given as the former case Fig.3. However, system can not learn, because closed circuit is not formed between STM and LTM.

Fig.3 shows this behavior. Closed circuit eventually dissolved because of the increase of threshold, which is the result of fatigue effect. At time T4, the system only receive main input just the same as time T1. However, output cell of LTM is activated because the signal is transmitted through strengthened connection W. These behavior shows that the system accomplished learning by associative input around T2 region. The system showed different behavior before learning and after learning. It is important to note that learning is processed with a help of closed circuit between long-term memory and short-term memory.

Next we examine how the system without short-term memory behaves for the same input series. Figure 4 shows this case. Although the input series are the same as Fig. 3, system can not learn, namely, no increase of weight W is found. This is due to the luck of short-term memory. We can observe this situation around T2 and T3 region in Fig.3. Output of LTM y is not firing at T4.

4. Discussion

We proposed the neural network model composed of long-term memory and short-term memory, which can be classified into a group of recurrent neural network model. These two area LTM and STM played different roles each ether in the total system. Long-term memory area is the place where memories can be stored as a variety of synaptic connections. These memories i.e. synaptic weight values were kept very stably in LTM area, because each synaptic weight has two stable points: W=0 no-connection and W=1 full-connection. This mechanizm was realized by introducing potential function of synaptic weight $\Phi(W)$: $\Phi(W) = -(W - \frac{1}{2})^2 + 2(W - \frac{1}{2})^4$.

We already discussed about validity of this postulate on physiological basis in the previous paper[5]. We built STM as the area where cells in it help to change the synaptic weight in LTM by forming closed circuit between them. We showed the simulation results that the system lost the ability of learning in the case of absence of short-term memory area. This behavior can naturally be understood, because closed circuit could not formed between LTM and STM. Simulation results are in principle in agreement with findings or hypotheses stated by many physiologists as we mentioned in the introduction.

Finally we discuss several problems which we have not mentioned so far.

1. We have not included the effect of well known phenomenon long-term potentiation (LTP) founded in hippocampus. Concerning to this point, we can state as follows. We introduced the fatigue effect as rising threshold value for eventually stopping firing of cells in closed circuit. However, we can not correctly point out at present how this effect contributes to the functional of association and learning. LTP phenomenon may has relation to this point,

and may help to form the formation of closed circuit.

2. We have not mentioned about the problem of time scale, namely, how the time factor appeared in the model relates with time scale in actual brain. Although almost mathematical neural network models have not directly discussed about this problem, it seems important to study about this for heading toward constructing hard architecture based on neural network theory in future. Including this point, it is said that the present work plays a role to bridge between physiological findings and artificial intelligent system.

In near future, parallel processing computer will surely be available more easier and cheaper than now. It is said that the study about the characteristics of neural network as shown in this article has an important role to construct artificial intelligent system.

References

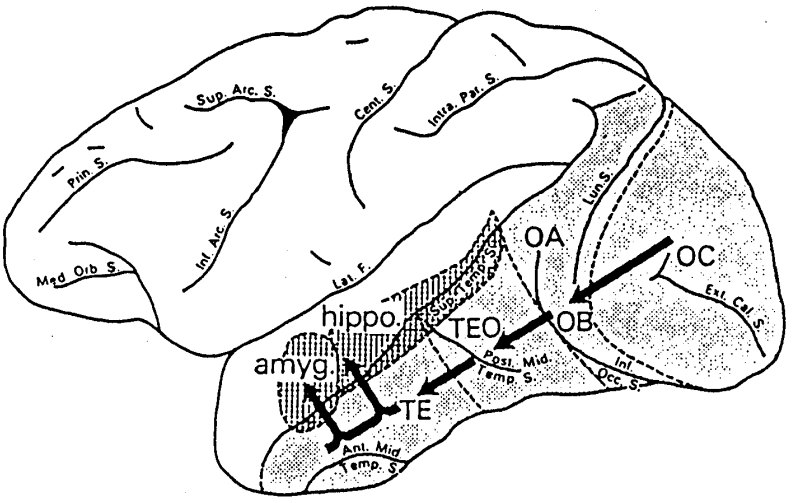
1) M.Mishikin: "A Memory System in the Monkey", Phil. Trans. Roy. Soc. Lond. **B298**(1982)85-95.

2) E.Iwai: *nō: gakusyu·kioku no mechanizm* (Brain: Mechanizm of learning and memory) (Asakura-Shoten, Tokyo, 1984) [in Japanese]
E.Iwai and M.Mishikin: "Two visual foci in the temporal lobe of monkeys. *in Neurophysiological basis of learning and behavior*, (ed. N.Yoshii and N.A. Bouchwald, Osaka, Osaka Univ.Press 1968).

3) S.Amari and M.Arbib: "Competition and cooperation in neural nets": *in Systems neuroscience*, (J.Metzler(ed.), MIT Press, 1977).

4) R.J.Williams and D.Zipper: " A learning algorithm for continually running fully recurrent neural networks", Neural Computation, **1**,(1989)643-644; B.A.Pearlmutter: "Learning state space trajectories in recurrent neural networks", Neural Computation, **1**(1989)263-269; M.Sato: "A learning algorithm to teach spatio-temporal patterns to recurrent neural networks",Biol.Cybern.,**62**(1990)259-263.

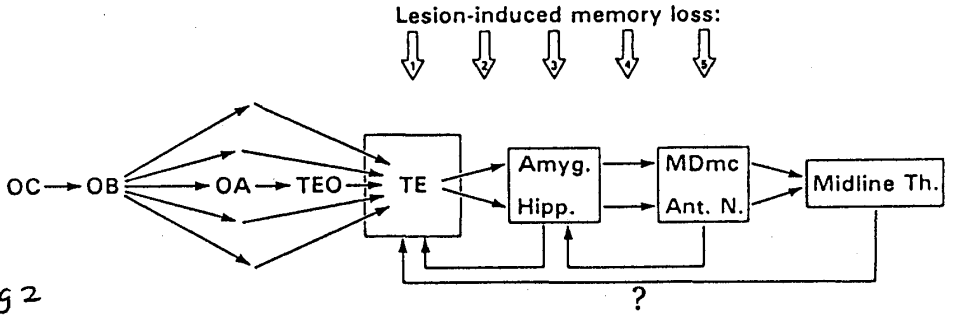
5) Y.Usami: "Pattern formation of synaptic connections in a generalized Hebb-type learning model", J.Phys.Soc.Jpn. **61**.(1992)735-743.



A. Fig 1

Flow of visual information from primary cortical area (OC) through secondary areas (OB, OA and TEO) to the highest-order visual area (TE), and from there into the medially located amygdaloid complex (amyg.) and hippocampal formation (hippo.). Cytoarchitectonic designations are those of Bonin & Bailey (1947). For clarity, the hippocampal formation is pictured slightly dorsal to its actual location.

M. Mishikin (1982)



A. Fig 2

The postulated circuit for visual recognition memory. Visual information is distributed from area OC for submodality processing within the prestriate complex (areas OB, OA, and TEO) and is then reintegrated in area TE. The convergent inputs to area TE are stored as central representations of stimuli, provided that area TE activates either an amygdalo-thalamic or hippocampo-thalamic pathway, which then feeds back to strengthen the prestriate–TE synapses either through reciprocal connections or via a relay in the midline thalamus. Severe visual recognition losses have been induced by lesions at each of the five indicated loci. See text for further explanation. MDmc, magnocellular portion of the dorsomedial nucleus; Ant. N., anterior nuclei.

M. Mishikin (1982)

Mishikin の提唱する視覚系記憶回路 II